Raghav Arora

+918295959517 rkinger2004@gmail GitHub/RaghavArora14 LinkedIn/RaghavArora14

Education

B.Tech Computer Science and Engineering (CGPA 7.41**)** Sep, 2022 - July,2026

BML Munjal University - Currently 3rd Year Gurgaon ,Haryana

- **Relevant Coursework:** Machine Learning, Computer Vision, Natural Language Processing, Big Data Analytics , OOPs and DSA (Python), Data Science, Computer Organization & Architecture, Software Engineering , DBMS, OS.

Machine Learning Specialization: Mar, 2024 - June,2024

DeepLearning.AI and Stanford Online

- **Relevant Coursework:** Supervised Learning (Regression and Classification), Advanced Algorithms, Neural Networks, Working with Tensorflow, Decision Trees, Unsupervised Learning (Clustering and Reinforcement Learning).

Skills

- **Languages**: Python, SQL, Shell, C++
- **Frameworks and Technologies**: Tensorflow, PyTorch, Flask, NumPy, Pandas, Git and Github, MySQL.
- **Concepts**: Machine Learning, Transformers , Neural Networks, Natural Language Processing, Computer Vision, Version Control, OOPs, Data Structures and Algorithms
- **Soft Skills**: Teamwork, Technical Presentation & Reports, Communication, Efficiency with resources

Hackathons

Contextual Vision and Language ChatBot | SIH-24 *Github*

- Led a team to get shortlisted in internal round of SIH'24 for the project titled "Contextual Vision Chatbot"
- Made a finetuned QWEN2VL based ChatBot for Vision and Language Processing.
- Managed to increase the token length of the conversation by 90% and decrease memory usage by 50% by making use of various memory saving techniques

Projects

Fine-Tuning MuRIL for Sentiment Analysis in Indian Languages | *MuRIL*, *Flask* ,*Tensorflow*

- Finetuned MuRIL (Multilingual Representation of Indian Languages) model for Sentiment analysis in Hindi and Punjabi Punjabi Language achieving upwards of 77% accuracy and F1 scores.
- Conducted a Comparative Analysis of MuRIL comparing it to various Multilingual BERT Models and other traditional methods to achieve upto 20% Improvement in Sentiment Analysis across both Languages
- Explored Various Hyperparameters for finetuning the model and which resulted in upwards of 75% mean validation accuracy while performing K-Fold Cross Validation

Legal Document Summarizer using LED | *LED*, *Torch* ,*Flask* *Github*

- Delivered concise, accurate summaries with enhanced clarity, significantly reducing processing time and increasing the input token limit by 700% and making legal content more accessible.
- Legal documents are lengthy and complex, with traditional models limited to processing 512-1024 tokens, often missing critical context.
- Fine-tuned an LED-based summarization system, increasing the token input limit to 8192. Utilized a dataset of 7,030 legal documents containing legal arguments and citations.

Raghav Arora

+918295959517 rkinger2004@gmail GitHub/Raghav Arora LinkedIn/Raghav Arora

Projects

SmartBin- Machine Learning based Smart Waste Classifier| *AWS, Flask, Tensorflow* *Github*

- Led the creation of a high-accuracy real-time image based waste classification system that could efficiently process and classify images while being deployed on resource-constrained edge devices.
- Implemented a CNN with 9.86M trainable parameters on AWS EC2 (t3.micro), using Flask for API endpoints, and integrated it with a Raspberry Pi edge device for real-time image classification and monitoring.
- Achieved 96.13% classification accuracy on biodegradable/non-biodegradable waste, with successful deployment of an end-to-end ML pipeline from edge capture to cloud inference and monitoring.

Web Server Log Analysis for a Poland Based E-Commerce Website| *BDA, Python, Streamlit* *Github*

- Analyzed 35 million unstructured web server logs to identify traffic patterns and user behavior.
- Built a Streamlit-based tool for preprocessing and visualizing metrics like HTTP methods, response codes, and user sessions, uncovering 14,402 peak hourly requests and dominant static resource usage.
- Delivered actionable insights, optimizing resource allocation during peak hours and improving user experience, particularly for mobile platforms.

Bi-Layer Segmentation of Post-Operative OCT Macular Hole Images| *UNET, VIA, PyTorch*

- Complex segmentation of ILM and RPE layers in post-operative OCT images required extensive manual annotation to create a reliable training dataset.
- Implemented a modified U-Net architecture with focal loss integration after manually annotating OCT images using VGG Image Annotator to label ILM and RPE layers.
- Accomplished 98% training and 97% test accuracy, successfully developing an automated system for retinal layer analysis that enables quantitative measurement of ILM-RPE distances.